



HAL
open science

Différences acoustiques inter-genres chez des bilingues Anglais/Français : une étude du Voice Onset Time

Erwan Pépiot, Aron Arnold

► **To cite this version:**

Erwan Pépiot, Aron Arnold. Différences acoustiques inter-genres chez des bilingues Anglais/Français : une étude du Voice Onset Time. XXXIVe Journées d'Études sur la Parole – JEP 2022, Université de Nantes, Jun 2022, Noirmoutier, France. pp.81-89, 10.21437/JEP.2022-9 . hal-03998837

HAL Id: hal-03998837

<https://hal.science/hal-03998837>

Submitted on 21 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Différences acoustiques inter-genres chez des bilingues Anglais/Français : une étude du Voice Onset Time

Erwan Pépiot¹, Aron Arnold²

(1) TransCrit (groupe LECSeL) – Université Paris 8, 2 rue de la liberté, 93526 Saint-Denis,
France

(2) VALIBEL – Université catholique de Louvain, Place Blaise Pascal 1, 1348 Louvain-la-
Neuve, Belgique

erwan.pepiot@free.fr, aron.arnold@uclouvain.be

RÉSUMÉ

Cette étude s'intéresse aux productions de locutrices et locuteurs bilingues anglais/français. Les Voice Onset Time (VOT) des consonnes occlusives sourdes et voisées ont été mesurés dans les deux langues, lors d'une tâche de lecture de mots. Les résultats indiquent un effet significatif des facteurs *langue* et *genre* sur ce paramètre. Le VOT des occlusives sourdes est plus long chez les locutrices dans les deux langues, avec une différence inter-genre accrue sur l'anglais. Les occlusives voisées présentent un pré-voisement plus long chez les femmes en français, et un VOT proche de zéro pour les deux genres en anglais. Le contraste de VOT entre les consonnes sourdes et leurs contreparties voisées est significativement plus long chez les locutrices dans les deux langues. Ce phénomène est légèrement plus marqué sur l'anglais. Ces données attestent l'existence de normes phonétiques socialement construites et dépendantes du genre et de la langue parlée, auxquelles se conforment les locutrices et locuteurs bilingues.

ABSTRACT

Cross-gender acoustic differences in English/French bilinguals: a study of Voice Onset Time.

The present study focuses on the productions of female and male English/French bilingual speakers. Voice Onset Time (VOT) of voiced and voiceless plosive consonants produced in a reading task was measured in both languages. Results show a significant effect of gender and language on this parameter. VOT of voiceless stops was longer in Female speakers in both languages, with a greater difference in English. French voiced plosives exhibit a longer pre-voicing in Female speakers, while the VOT of English voiced plosives is close to zero for both genders. VOT contrast between voiceless stops and their voiced counterparts was also significantly longer in female speakers in both languages. The scope of this cross-gender difference was greater in English. This resulted in a smaller cross-gender difference in French for this parameter. These data show the existence of socially constructed language- and gender-specific phonetic norms, to which bilingual speakers seem to adapt.

MOTS-CLÉS : voix et genre, différences inter-genres, VOT, Voice Onset Time, consonnes occlusives, français parisien, anglais américain, bilinguisme, variations inter-langues.

KEYWORDS: voice and gender, cross-gender differences, VOT, Voice Onset Time, plosive consonants, Parisian French, American English, bilingualism, cross-language variation.

1 Introduction

Les différences acoustiques entre les voix de femmes et d'hommes ont fait l'objet d'un nombre important de recherches depuis les années 1970. Ces études portent en majorité sur la fréquence fondamentale (F0) et les fréquences de résonance (en particulier les formants vocaliques), souvent présentées comme les deux principaux paramètres permettant qu'une voix soit perçue comme féminine ou masculine.

Le F0 des voix féminines se trouve en moyenne dans des fréquences plus élevées que celui des voix masculines (Pépiot 2014a), quant aux formants vocaliques (fréquences de résonance des voyelles), ils sont généralement situés dans des fréquences plus basses chez les locuteurs masculins (Pépiot, 2016a). L'origine de ces variations acoustiques est traditionnellement attribuée aux différences femmes/hommes qui émergent lors de la puberté dans les appareils phonatoires, avec un allongement du conduit vocal et des plis vocaux en moyenne plus marqué chez les locuteurs masculins (Kahane, 1978). Une explication remise en partie en cause par plusieurs études mettant en avant la dimension socialement construite de ces différences vocales (Arnold, 2015 ; Pépiot 2013, 2014b).

Plusieurs auteur·rices ont également étudié les différences inter-genres sur le *Voice Onset Time* (VOT) des consonnes occlusives. Ce paramètre, qui correspond au délai entre le relâchement d'une consonne et l'entrée en vibration des plis vocaux, varie en fonction de la nature de la consonne : un VOT relativement long sur les consonnes sourdes, plus élevé encore lorsqu'elles sont « aspirées », et proche de zéro, voire négatif, sur les consonnes voisées. Concernant les variations femmes/hommes, chez les locuteur·ices anglophones, la tendance générale est à l'observation d'un VOT plus long chez les femmes sur les occlusives sourdes (Swartz, 1992 ; Whiteside & Irving, 1997 ; Robb & al., 2005). Dans les autres langues, les résultats sont nettement moins clairs. En Coréen, par exemple, les locuteurs produiraient des VOT équivalents à ceux des femmes, voire plus courts dans certains cas (Oh, 2011). De même, une recherche menée sur des locuteur·ice·s suédois·e·s ne montre aucune différence inter-genre significative (Karlsson & al., 2004). Sur le Français, une étude de Pépiot (2016) indique que les femmes produisent des VOT plus longs que les hommes sur les occlusives sourdes et plus courts sur les occlusives voisées (i.e. pré-voisement plus important).

Robb et al. (2005), avancent l'hypothèse selon laquelle la morphologie des plis vocaux des femmes serait responsable d'une pression sub-glottale plus importante que chez les hommes lors de la phase de fermeture des plis vocaux, entraînant ainsi un démarrage du voisement plus tardif (c'est-à-dire un VOT supérieur). Koenig (2000) propose une explication similaire : les plis vocaux des femmes seraient intrinsèquement plus tendus que ceux des hommes. Ainsi, les plis vocaux des hommes, plus relâchés, pourraient entrer en vibration plus rapidement (VOT plus court). Mais cette explication vient se heurter aux résultats exposés plus haut. Comment expliquer ces variations inter-langues, avec des VOT plus courts chez les femmes en coréen et sur les occlusives voisées du français ?

Le croisement de ces différentes recherches suggère donc que ces variations vocales inter-genres seraient en partie culturelles et construites socialement, car dépendantes de la langue parlée. Se pose alors la question des locuteur·ices bilingues. Comment s'adaptent-elles/ils aux normes genrées des différentes langues parlées ? Ces questions n'ont pour l'instant fait l'objet que d'une attention assez limitée dans les sciences phonétiques. S'il existe un certain nombre d'études ayant comparé les productions de locuteur·ices de différentes langues (voir notamment Pépiot 2014a, 2016b), peu de chercheur·euses se sont penché·es sur les variations intra-individuelles des productions vocales des bilingues lors du passage d'une langue à l'autre et ont investigué l'influence éventuelle du facteur « genre » dans ces différences.

Les quelques recherches conduites sur des locuteur·ices bilingues et ont montré que ces dernier·es allaient faire varier leur F0 moyen en fonction de la langue parlée (Altenberg & Ferrand, 2006, sur des bilingues russe/anglais ; Lee, Van Lanker Sidtis, 2017, sur des bilingues coréen/mandarin). Ces analyses ont cependant été effectuées uniquement sur des productions de locutrices – il est donc impossible de savoir si les variations observées relèvent d’une adaptation aux normes genrées des différentes langues, ou tout simplement de pratiques liées aux langues elles-mêmes, sans considération de genre. D’autres études plus récentes, menées quant à elles sur des bilingues femmes et hommes, tendent à préciser ce phénomène. Ainsi, chez les bilingues anglais/français, les locutrices auraient tendance à utiliser un F0 moyen nettement plus élevé en français qu’en anglais, ce qui n’est pas le cas des locuteurs masculins (Pépiot & Arnold, 2018). Sur cette même population, des variations genrées différentes selon la langue parlée ont également été observées sur les formants vocaliques et la qualité de voix (Pépiot & Arnold, 2020).

Pour l’heure, ces recherches se sont cependant limitées aux paramètres évoqués précédemment et n’ont pas investigué le Voice Onset Time, un paramètre acoustique pourtant crucial et faisant l’objet d’un intérêt certain dans la littérature phonétique depuis plusieurs décennies. De plus, un travail portant sur des bilingues anglais/français nous paraît particulièrement intéressant dans la mesure où ces deux langues présentent des caractéristiques de VOT très différentes : un VOT légèrement positif sur les occlusives sourdes et nettement négatif sur les consonnes voisées en français, un VOT fortement positif sur les occlusives sourdes (dites *aspirées*) et proche de zéro sur les occlusives voisées en anglais (Pépiot, 2016b). Nous avons donc souhaité nous pencher sur le VOT des locutrices et de locuteurs bilingues anglais L1 / français L2 avec l’hypothèse de départ suivante : *les locutrices et les locuteurs bilingues adaptent leurs pratiques vocales aux normes genrées de la langue employée, y compris pour le VOT des consonnes occlusives*. Nous soutenons que l’étude de ces variations inter-langues dans les productions des bilingues permet de s’éloigner d’une vision statique des voix de femmes et d’hommes, dans laquelle celles-ci sont présentées comme une caractéristique essentielle des locuteur·ices et comme étant uniquement liées à leur anatomie, pour en adopter une vision plus dynamique permettant d’intégrer des questions de différences culturelles, de normes et de performances de genre.

2 Méthode

2.1 Corpus

Notre étude se base sur l’analyse d’un corpus de mots dissyllabiques en anglais et en français collectés lors d’une tâche de lecture :

- Combinaisons /C (occlusive) – V – p – i/: /pipi/, /papi/, /pupi/, /bipi/, /bapi/, /bupi/, /tipi/, /tapi/, /tupi/, /dipi/, /dapi/, /dupi/, /kipi/, /kapi/, /kupi/, /gipi/, /gapi/, /gupi/ pour le corpus français, /pi:pi/, /pæpi/, /pu:pi/, /bi:pi/, /bæpi/, /bu:pi/, /'ti:pi/, /'tæpi/, /'tu:pi/, /'di:pi/, /'dæpi/, /'du:pi/, /'ki:pi/, /'kæpi/, /'ku:pi/, /'gi:pi/, /'gæpi/, /'gu:pi/ pour le corpus anglais.
- Combinaisons /C (fricative) – V – p – i/: /sipi/, /sapi/, /supi/, /zipi/, /zapi/, /zupi/, /ʃipi/, /ʃapi/, /ʃupi/, /zipi/, /zapi/, /zupi/ pour le corpus français, /'si:pi/, /'sæpi/, /'su:pi/, /'zi:pi/, /'zæpi/, /'zu:pi/, /'ʃi:pi/, /'ʃæpi/, /'ʃu:pi/, /'zi:pi/, /'zæpi/, /'zu:pi/ pour le corpus anglais.
- Combinaisons /V – p – i/: /ipi/, /api/, /upi/ pour le corpus français, /'i:pi/, /'æpi/, /'u:pi/ pour le corpus anglais.

Il n’existe pas d’accent lexical en Français parisien d’un point de vue phonologique, mais au sein de la phrase porteuse utilisée ici (voir 2.3) les locuteur·ices ont systématiquement accentué la première syllabe des mots du corpus.

2.2 Participant·e·s

Six locuteurs et six locutrices bilingues anglais américain L1 / français parisien L2 ont pris part à cette étude. Les participant·e·s sont originaires du nord-est des États-Unis et vivent en région parisienne depuis plusieurs années. Tou·te·s font état d'une pratique quotidienne de la langue française et d'un niveau d'aisance supérieur ou égal à 3 sur une échelle allant de 0 à 5 via un questionnaire inspiré de celui établi par Grosjean (2013). Nous reprenons également à notre compte la définition du bilinguisme proposé par cet auteur.

Les participant·e·s étaient âgé·e·s de 29 à 54 ans ($SD=7,6$ ans) au moment des sessions d'enregistrement, avec une moyenne d'âge de 41,8 ans chez les femmes et de 40 ans chez les hommes. Aucun·e n'était fumeur·euse ou ne présentait de trouble de la parole. Une clé USB a été offerte en échange de la participation à la présente étude.

2.3 Procédure d'enregistrement

Les enregistrements se sont déroulés dans une chambre anéchoïque, en utilisant un enregistreur numérique *Edirol R09-HR* de marque *Roland*. Il a été demandé aux participant·e·s de lire les mots dissyllabiques listés dans la section 2.1. Ces derniers ont été présentés sous leur forme orthographique et placés dans une phrase porteuse afin de maintenir les paramètres prosodiques constants : “He said ‘*MOT DISYLLABIQUE*’ twice” pour le corpus anglais et “Il a dit ‘*MOT DISYLLABIQUE*’ deux fois” pour le corpus français. Les locuteur·ices ont prononcé chacune des phrases deux fois.

Afin de neutraliser de possibles biais liés à l'ordre d'emploi des deux langues (voir Altenberg, Ferrand, 2006), la moitié des participant·e·s a démarré par la lecture du corpus anglais, et l'autre moitié par celui en langue française.

2.4 Analyse des données

L'analyse acoustique du corpus a été effectuée à l'aide du logiciel *Praat*. Une fois les mots dissyllabiques extraits des phrases porteuses, le Voice Onset Time (en ms) des consonnes occlusives initiales a été mesuré (33 mots * 2 répétitions par locuteur·ice). Cette mesure correspond à l'intervalle entre le relâchement de la consonne et le début du voisement. Si celui-ci démarre avant le relâchement, ce qui est généralement le cas sur les occlusives voisées du français (Pépiot, 2016b), la valeur est alors négative. Ces données ont été mesurées manuellement à l'aide des spectrogrammes, en repérant visuellement le relâchement de la consonne et le début de la vibration des plis vocaux.

3 Résultats

3.1 VOT moyens

Les VOT moyens (en ms) des consonnes occlusives du français sont présentés dans le Tableau 1 (6 occurrences par consonne et par locuteur·ice).

Sur les consonnes sourdes, le VOT est en moyenne plus élevé chez les femmes et ce, quelle que soit la consonne ([p], [t] et [k]). Sur l'ensemble des consonnes, le VOT est en moyenne 21% plus élevé chez les locutrices (+9ms). On remarque également que, indépendamment du genre des locuteur·ices, les valeurs obtenues suivent les tendances universelles du VOT : [p] < [t] < [k].

VOT (MS) DES CONSONNES FRANÇAISES

Loc.	[p]	[t]	[k]	Moy. cons. sourdes	[b]	[d]	[g]	Moy. cons. voisées
F1	58	64	89	70	-89	-69	-91	-83
F2	16	32	41	30	-73	-72	-72	-72
F3	53	64	64	60	-53	-34	-89	-58
F4	44	58	61	55	-26	-70	-40	-45
F5	44	63	70	59	-79	-83	-83	-82
F6	17	40	56	38	-84	-74	-84	-81
Moy. F.	39	53	64	52	-67	-67	-76	-70
H1	24	38	42	34	-115	-91	-91	-99
H2	34	57	57	50	-85	-59	-64	-70
H3	36	48	59	47	-46	-38	-34	-39
H4	15	32	43	30	-80	-49	-35	-55
H5	44	56	58	53	-52	-51	-58	-53
H6	36	46	53	45	-71	-86	-26	-61
Moy. H.	31	46	52	43	-75	-62	-51	-63

TABLE 1 : VOT moyen (ms) des locutrices (F) et les locuteurs (H) sur les mots dissyllabiques du corpus français (6 occurrences par consonne et par locuteur·ice).

Les consonnes voisées présentent toutes des valeurs négatives en raison du pré-voisement, phénomène caractéristique des occlusives voisées du français (Pépiot, 2016b). En moyenne, ce pré-voisement est plus long chez les locutrices, avec 7 millisecondes supplémentaires par rapport aux locuteurs (+11%).

Une ANOVA à deux facteurs (« *genre* » & « *consonne* ») a été effectuée sur les VOT des occlusives sourdes du français. Ce test révèle un effet significatif du facteur *genre* ($F(1,210)=12,120$; $p<0,001$) et du facteur *consonne* ($F(2,210)=27,981$; $p<0,0001$). Aucune interaction significative n'apparaît entre les deux facteurs, ce qui confirme que les VOT des locutrices sur les occlusives sourdes tendent à être significativement plus longs, indépendamment de la consonne considérée.

Le même test été conduit sur les occlusives voisées du français. Malgré une différence inter-genres notable, l'influence du facteur *genre* s'avère non significative, probablement en raison du trop faible nombre d'occurrences : $F(1,210)=1,869$; $p=0,1730$. Le facteur consonne n'a pas non plus d'influence significative ($F(2,210)=0,683$; $p=0,5064$).

Les VOT moyens mesurés dans les mots anglais dissyllabiques sont présentés dans le tableau 2, ci-après (6 occurrences par consonne et par locuteur·ice).

Tout comme en français, les VOT des consonnes sourdes sont plus courts chez les locuteurs que chez les locutrices, quel que soit la consonne. Cette tendance se trouve même accentuée en anglais, avec des VOT de consonnes sourdes en moyenne 28% plus élevés chez les locutrices (+16ms). Ces valeurs suivent également les tendances dites universelles des VOT, dans lesquelles locutrices et locuteurs se confondent : $[p] < [t] < [k]$.

Sur les occlusives voisées, les femmes et les hommes présentent des valeurs relativement similaires. La valeur moyenne, toutes consonnes confondues, se situe autour de 0 milliseconde pour les deux genres étudiés, c'est-à-dire une mise en vibration des plis vocaux concomitante avec le relâchement de la consonne. Ces données sont très cohérentes avec les résultats généralement obtenus chez des locuteur·ices monolingues, sauf pour *H1* et *F2* : ces deux participant·es ont systématiquement produit un « pré-voisement », ce qui semble clairement indiquer une interférence de leur L2 (le français) sur leur L1 (l'anglais).

VOT (MS) DES CONSONNES ANGLAISES

Loc.	[p ^h]	[t ^h]	[k ^h]	Moy. cons. sourdes	[b]	[d]	[g]	Moy. cons. voisées
F1	81	93	116	96	11	14	29	18
F2	32	58	62	51	-83	-92	-88	-88
F3	65	76	81	74	10	16	25	17
F4	46	65	77	63	11	20	24	18
F5	65	80	87	77	7	20	34	20
F6	62	80	100	81	7	17	23	15
Moy. F.	58	75	87	74	-6	-1	8	0
H1	42	61	57	53	-46	-14	-59	-40
H2	40	56	64	53	6	19	24	16
H3	47	64	71	61	11	21	28	20
H4	45	50	71	55	11	15	26	17
H5	45	65	56	55	-16	-35	15	-12
H6	60	67	78	68	-7	11	9	4
Moy. H.	46	61	66	58	-7	3	7	1

TABLE 2 : VOT moyen (ms) des locutrices (F) et les locuteurs (H) sur les mots dissyllabiques du corpus anglais (6 occurrences par consonne et par locuteur·ice).

Une ANOVA à deux facteurs (« *genre* » et « *consonne* ») a été effectuée pour les valeurs du VOT des occlusives sourdes de l'anglais. Cette analyse montre un effet large et significatif du facteur *genre* ($F(1,210)=49,491$; $p<0,0001$) ainsi que du facteur *consonne* ($F(2,210)=39,777$; $p<0,0001$). Aucune interaction significative n'a été trouvée entre ces deux facteurs. Cela confirme que, de manière similaire aux faits observés sur le français, les VOT produits par les locutrices tendent à être significativement plus longs que ceux produits par les locuteurs masculins lors de la production d'occlusives sourdes, indépendamment de la consonne.

Le même test statistique a été conduit sur le VOT des occlusives voisées de l'anglais. Sans surprise, cette ANOVA indique qu'il n'y a pas d'effet significatif du facteur *genre* sur ces valeurs ($F(1,210)=0,030$; $p=0,8628$). De plus, aucun effet significatif du facteur *consonne* n'a été trouvé ($F(2,210)=2,415$; $p=0,0919$), ni une interaction entre ces deux facteurs.

3.2 Contraste de VOT (occlusives sourdes vs. voisées)

Le contraste de VOT, qui correspond à la différence entre le VOT d'une occlusive sourde et le VOT de sa contrepartie voisée, est présenté dans le tableau 3, pour le français et l'anglais.

De manière générale, le contraste de VOT est nettement supérieur dans les productions des locutrices par rapport à celles des locuteurs. Ce phénomène s'observe dans les deux langues, mais est plus marqué en anglais, avec un contraste de VOT en moyenne 30% supérieur chez les femmes (alors que celui-ci n'est que de 15% en français).

Afin de tester si ces différences inter-genres sont significatives, une ANOVA à deux facteurs (« *genre* » et « *type de consonne* ») a été réalisée sur ces données. En français, il existe un effet significatif du genre des locuteur·ices ($F(1,102)=5,173$; $p<0,01$), mais pas d'effet du type de consonne ($F(2,102)=1,568$; $p=0,2134$) ni d'interaction entre ces deux facteurs.

Loc.	CONTRASTE DE VOT EN FRANÇAIS (MS)				CONTRASTE DE VOT EN ANGLAIS (MS)			
	[p] vs. [b]	[t] vs. [d]	[k] vs. [g]	Sourdes vs. voisées	[p ^h]vs. [b]	[t ^h] vs. [d]	[k ^h] vs. [g]	Sourdes vs. voisées
F1	147	132	180	153	70	79	87	78
F2	89	104	113	102	116	150	150	138
F3	106	98	152	119	55	60	57	57
F4	70	128	101	100	35	45	53	44
F5	124	146	154	141	58	61	54	57
F6	101	115	140	119	56	64	78	66
Moy. F.	106	120	140	122	65	76	80	74
H1	138	128	133	133	87	75	117	93
H2	120	116	121	119	34	37	41	37
H3	81	86	93	87	37	44	43	41
H4	95	82	78	85	35	36	45	38
H5	95	107	116	106	60	99	41	67
H6	107	132	79	106	67	57	69	64
Moy. H.	106	109	103	106	53	58	59	57

TABLE 3 : Contraste de VOT moyen (ms) entre les occlusives sourdes et voisées chez les locutrices (F) et les locuteurs (H) bilingues sur des mots dissyllabiques, en fonction de la langue parlée (français ou anglais).

Des résultats similaires ont été trouvés sur les consonnes anglaises, avec un effet significatif du facteur *genre* ($F(1,102)=7,159$; $p<0.01$), une absence d'effet significatif pour le type de consonne ($F(2,102)=1,035$; $p=0,3588$), et aucune interaction entre ces facteurs.

Cela indique que, dans les deux langues, le contraste de VOT tend à être significativement plus long dans les productions féminines, et ce indépendamment du lieu d'articulation de la consonne. Tout semble ainsi indiquer une variation d'ordre sociophonétique ayant pour origine des conduites articulatoires genrées.

4 Conclusion / discussion

Cette analyse inter-langues a permis d'obtenir plusieurs résultats intéressants, qu'il convient d'interpréter et de mettre en perspective. Globalement, nos résultats montrent que le VOT des consonnes occlusives sourdes est significativement plus long chez les locutrices que chez les locuteurs masculins. Cela s'est vérifié à la fois en anglais et en français : les valeurs des locutrices sont 21% plus élevées que celles des locuteurs en français, et 28% plus élevées en anglais. Ces résultats sont en adéquation avec ceux observés chez des anglophones monolingues (Pépiot 2016b ; Robb & al., 2005 ; Swartz, 1992 ; Whiteside & Irving, 1997) et chez des francophones monolingues (Pépiot, 2016b). Indépendamment du genre des locuteur·ices, on note que ces valeurs suivent les tendances universelles du VOT observés dans la littérature ($[p] < [t] < [k]$).

Sur les occlusives voisées, les résultats sont très similaires au sein des deux catégories de genre que nous avons étudiées en anglais, avec des valeurs proches de zéro, c'est-à-dire une quasi-concomitance entre le relâchement de la consonne et le début du voisement. Cela est là aussi en adéquation avec les données observées chez les anglophones monolingues (Swartz, 1992, Pépiot 2016b). Par ailleurs, deux participant·es – un homme et une femme – ont produit un important pré-voisement sur les occlusives voisées de l'anglais, ce qui constitue un fait notable. Il s'agit très vraisemblablement de l'influence de leur langue seconde (le français, devenu dominant) sur leur langue première (l'anglais), ce type de phénomène est largement documenté dans les recherches sur le bilinguisme (voir notamment Ryst & Pépiot, 2015). En français, bien que cette différence

n'atteigne pas le seuil de significativité statistique, les locutrices présentent un pré-voisement nettement plus précoce que les locuteurs (+11% en moyenne). Cela recoupe les données obtenues par Pépiot (2016b) sur des francophones monolingues.

Le contraste de VOT, qui correspond à la différence entre le VOT d'une occlusive sourde et le VOT de sa contrepartie voisée, a aussi été calculé. En anglais comme en français, ce contraste est significativement plus long chez les locutrices, avec une différence proportionnellement plus grande en anglais (+30% par rapport aux locuteurs) qu'en français (+15%), et ce malgré le fait que le renforcement de ce contraste en anglais ne passe que par l'allongement du VOT des consonnes sourdes (peu de marge de variation pour les consonnes dites voisées qui se caractérisent intrinsèquement par un VOT proche de 0 dans cette langue).

Ce phénomène peut être vu comme l'illustration d'une tendance sociophonétique plus large : la production d'un plus grand contraste de VOT peut être en effet considérée comme un moyen pour les locutrices de renforcer l'intelligibilité de leur parole. Cette tendance à une articulation plus « soignée » chez les femmes est observée dans de nombreux paramètres acoustiques, tels que la dispersion des formants vocaliques, la répartition temporelle consonnes/voyelles ou encore la modulation de F0 (voir Pépiot, 2014b, 2016a). Cela s'inscrit dans des stéréotypes de genre classiques de nos sociétés patriarcales, avec une injonction faite aux femmes de prêter plus attention à leur apparence, leur posture, leur manière d'être, par opposition aux hommes pour lesquels une attitude et/ou une apparence plus « négligée » est plus facilement admise socialement.

L'étude croisée des facteurs *langue* et *genre* dans le VOT des consonnes occlusives produites par des bilingues apporte ainsi de nouveaux éléments qui vont dans le sens des récentes observations faites par Pépiot et Arnold (2018, 2020) sur d'autres paramètres acoustiques, mais avec une méthodologie similaire. La présente étude confirme notamment que la production d'une voix de femme ou d'une voix d'homme mobilise des pratiques vocales différentes en fonction de la langue parlée, y compris au plan du VOT. Les différences inter-genres sur le VOT ne peuvent donc s'expliquer par des différences morphologiques femmes/hommes qui entraîneraient la mise en vibration plus rapide des plis vocaux chez les hommes (hypothèse avancée notamment par Koenig, 2000 et Robb & al., 2005), mais elles semblent bien résulter d'un apprentissage et d'une socialisation en tant que membre d'une catégorie de genre spécifique. Ceci constitue un nouvel argument pour s'éloigner des conceptions purement anatomistes de la voix, que l'on retrouve encore régulièrement dans la littérature phonétique, et pour prendre davantage en considération des facteurs sociaux et culturels.

Parmi les limites de cette étude, nous citerons notamment le nombre relativement réduit de locuteur·rices. Afin de confirmer ces tendances, la présente étude pourrait être répliquée avec un nombre plus important de participant·e·s. Par ailleurs, nous avons ici uniquement analysé de la parole produite par des bilingues anglais L1 / français L2. Il serait intéressant d'étudier aussi la parole de bilingues français L1 / anglais L2, mais aussi de bilingues d'autres langues. D'autre part, une étude menée sur des enfants bilingues pré-pubères, et ne présentant donc pas de différences morphologiques inter-genres au niveau du larynx, pourrait également apporter des éléments scientifiques particulièrement pertinents.

Remerciements

Nous adressons nous chaleureux remerciements aux locutrices et aux locuteurs ayant participé à cette étude, ainsi qu'à Juliette Stockman qui a contribué à la préparation de l'analyse de nos données.

Références

- ALTENBERG E. P., FERRAND C. T. (2006). Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. *Journal of Voice* 20, 89-96.
- ARNOLD A. (2015). Voix et transidentité : changer de voix pour changer de genre ?. *Langage et société* 151(1), 87-105.
- GROSJEAN, F. & LI, P. (2013). *The Psycholinguistics of Bilingualism*. Oxford : Wiley-Blackwell.
- KAHANE J. C. 1978. A morphological study of the human prepubertal and pubertal larynx. *American Journal of Anatomy* 151, 11-19.
- KARLSSON, F., ZETTERHOLM, E. & SULLIVAN, K. (2004). Development of a gender difference in voice onset time. In *Proceedings of the 10th Australian International Conference on Speech Science & Technology*, Sydney, 316-321.
- KOENIG, L. (2000). Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *Journal of Speech, Language and Hearing Research* 43, 1211-1228.
- LEE B., VAN LANCKER SIDTIS D. (2017). The bilingual voice: Vocal characteristics when speaking two languages across speech tasks. *Speech, Language and Hearing* 20(3), 174-185.
- OH, E. (2011). Effects of speaker gender on VOT in Korean stops. *Journal of Phonetics* 39, 59-67.
- PEPIOT, E. (2013). *Voix de femmes, voix d'hommes : différences acoustiques, identification du genre par la voix et implications psycholinguistiques chez les locuteurs anglophones et francophones*. Thèse de Doctorat. Université Paris VIII Vincennes-Saint Denis.
- PEPIOT E. (2014a). Male and female speech: a study of mean F0, F0 range, phonation type and speech rate in Parisian French and American English speakers. *Proceedings of the 7th International Conference on Speech Prosody*, 305-309.
- PEPIOT, E. (2014b). Voix et genre : un état de la question. In Ibrahim, A.H. (éd.), *La langue, la voix, la parole* (pp. 53-86), Paris : CRL.
- PEPIOT, E. (2016a). Female and male speech: a study of vowel formants and consonant noise in Parisian French and American English speakers. *Paisii Hilendarski University of Plovdiv, Bulgaria – Research Papers* 53(1), 610-621.
- PEPIOT, E. (2016b). Voix de femmes, voix d'hommes : une étude du voice onset time, de la répartition consonnes/voyelles et du débit de parole chez des locuteurs francophones et anglophones américains. *Actes de la conférence JEP-TALN-RECITAL 2016*, Vol.1, 759-767.
- PEPIOT, E. & ARNOLD, A. (2018). Étude des variations de fréquence fondamentale relatives au genre chez des bilingues Anglais/Français. *Actes de la conférence JEP 2018*, 178-186.
- PEPIOT, E. & ARNOLD, A. (2020). Différences acoustiques inter-genres chez des bilingues Anglais/Français : une étude des formants vocaliques et de la qualité de voix. *Actes de la conférence JEP-TALN-RECITAL 2020*, Vol.1, 480-488.
- ROBB, M., GILBERT, H. & LERMAN, J. (2005). Influence of gender and environmental setting on voice onset time. *Folia Phoniatica et Logopaedica* 57, 125-133.
- RYST, E. & PEPIOT, E. (2015). How do Parisian French and American English speakers prefer parsing CVCV words? A perceptual experiment to test the effects of linguistic context and L2 exposure. In *The Notion of Syllable Across History, Theories and Analysis* (pp. 326-353), Cambridge Scholars Publishing.
- SWARTZ, B. L. (1992). Gender difference in VOT. *Perceptual and Motor Skills* 75, 983-992.
- WHITESIDE, S. P. & IRVING, C. J. (1997). Speakers' sex differences in voice onset time, some preliminary findings. *Perceptual and Motor Skills* 85, 459-463.